UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/817,591	03/26/2001	Yihong Gong	CA1122	7751

23493 7590 07/07/2006

SUGHRUE MION, PLLC
401 Castro Street, Ste 220
Mountain View, CA 94041-2007

EXAMINER
TRAN, QUOC A

ART UNIT	PAPER NUMBER
2176	

DATE MAILED: 07/07/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/817,591	GONG ET AL.	
	Examiner	Art Unit	
	Quoc A. Tran	2176	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *11 April 2006*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-32* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-32* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

1. This action is responsive to Remarks filed 04/11/2006 with recognition of an original filing date of 03/26/2001, which claims benefit of 60/254,535 filed 12/12/2000.

2. Claims 1-32 are pending. Claims 1, 9, 13, 21, 26 and 29 are independent claims.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> ***(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.***

4. **Claims 1-20** are rejected under 35 U.S.C. 103(a) as being unpatentable by Billheimer et al. US Patent No. 6,611,825 B1 - filed 06/09/1999 (hereinafter Billheimer '825), in view of Goldstein et al. " Summarizing Text Documents: Sentence Selection and Evaluation Metrics" published 08/1999 by ACM Press (hereinafter Goldstein).

In regard to independent claim 1, creating a weighted document term-frequency vector for said document, (Billheimer '825 at col. 4, line 35 through col. 6, line 40, discloses an information retrieval method, wherein term and document visualization, term and document clustering, term and document classification, summarization of individual documents and groups of documents, and document cross-referencing. This is accomplished by representing the text of a document collection using subspace transformations. This subspace transformation representation is performed by: constructing a term frequency matrix of the term frequencies for

each of the documents, transforming the term frequencies for statistical purposes, and projecting the documents or the terms into a lower dimensional subspace).

Billheimer '825 does not explicitly teach, **for each sentence in said document, creating a weighted sentence term-frequency vector,** however (Goldstein at pages 121-122, discloses the method of generating summaries by text extraction, wherein scoring sentences with respect to both statistical and linguistic features such that a centroid query vector is calculated using high frequency document words and the title of the document wherein each sentence is scored accordingly (e.g. using Score formula on page 122) and then ordered in a summary according to rank order) Examiner read the above in the broadest reasonable interpretation to the claim limitation, wherein a weighted sentence term-frequency vector would have been an obvious variant of scoring sentences with respect to both statistical and linguistic features such that a centroid query vector is calculated using high frequency document words and the title of the document wherein each sentence is scored accordingly (e.g. using Score formula on page 122) and then ordered in a summary according to rank order, to a person of ordinary skill in the art at the time the invention was made.

computing a score for each said weighted sentence term-frequency vector in accordance with relevance to said weighted document, however (Goldstein at pages 121-122, discloses the method of generating summaries by text extraction, wherein each sentence is scored according to the following formula and then ordered in a summary according to rank order,

$$Score(S_i) = \lambda \sum_{s \in S} w_s * (Q_s \cdot S_i) + (1 - \lambda) * \sum_{l \in L} w_l * (L_l \cdot S_i)$$

where S is the set of statistical features, L is the set of linguistic features, Q is the query, and w is the weights for the features in that set. These weights can be tuned according to the type of data set used and the type of summary desired) Examiner read the above in the broadest reasonable interpretation to the claim limitation, wherein a weighted sentence term-frequency vector and accordance with relevance to said weighted document would have been an obvious variant of scoring sentences with respect to both statistical and linguistic features such that a centroid query vector is calculated using high frequency document words and the title of the document wherein each sentence is scored accordingly (e.g. using Score formula on page 122) and then ordered in a summary according to rank order, to a person of ordinary skill in the art at the time the invention was made.

selecting a sentence for inclusion in said generic text summary in accordance with said computing, however (Goldstein at pages 121-122, discloses the method of generating summaries by text extraction, wherein each sentence is scored according to the following formula and then ordered in a summary according to rank order.

$$Score(S_i) = \lambda \sum_{s \in S} w_s * (Q_s \cdot S_i) + (1 - \lambda) * \sum_{l \in L} w_l * (L_l \cdot S_i)$$

where S is the set of statistical features, L is the set of linguistic features, Q is the query, and w is the weights for the features in that set. These weights can be tuned according to the type of data set used and the type of summary desired. The evaluators also marked each document as relevant or not relevant to the topic and selected the three most relevant sentences for each article from the sentences that they had marked relevant (yielding a most relevant sentence data set of 1-9 sentences per document). This set has an average of 5.6 sentences per document and

58.2% of the relevant sentence summaries contain the first sentence).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Billheimer '825 teaching, wherein creating a weighted document term-frequency vector for said document, to includes a means of creating a weighted sentence term-frequency vector for each sentence in said document by computing a score for each said weighted sentence term-frequency vector in accordance with relevance to said weighted document, and selecting a sentence for inclusion in said generic text summary in accordance with said computing of Goldstein's teaching. One of ordinary skill in the art would have been motivated to perform such a modification to produce a query-relevant summary base upon the a weighted sentence term-frequency vector and accordance with relevance to said weighted document (Goldstein at pages 121-122).

In regard to independent claim 9, incorporate substantially similar subject matter as cited in claim 1 above, and further view of the following, and is similarly rejected along the same rationale,

a computer; a display for displaying said generic text summary; and summarizer program code, operable on said computer (as taught by Billheimer '825 at col. 8 line 25 through col. 9, line 5, also see Fig. 2),

for analyzing and summarizing said document; said summarizer program code (as taught by Billheimer '825 at col. 7 line 25 through col. 18, line 35).

Billheimer '825 does not explicitly teach, **a vector generator for creating a weighted document term-frequency vector for said document,** however (Goldstein at pages 121-122, discloses the method of generating summaries by text extraction, wherein scoring sentences with

respect to both statistical and linguistic features such that a centroid query vector is calculated using high frequency document words and the title of the document wherein each sentence is scored accordingly (e.g. using Score formula on page 122) and then ordered in a summary according to rank order, wherein each sentence is scored according to the following formula and then ordered in a summary according to rank order,

$$Score(S_i) = \lambda \sum_{s \in S} w_s * (Q_s \cdot S_i) + (1 - \lambda) * \sum_{l \in L} w_l * (L_l \cdot S_i)$$

where S is the set of statistical features, L is the set of linguistic features, Q is the query, and w is the weights for the features in that set. These weights can be tuned according to the type of data set used and the type of summary desired) Examiner read the above in the broadest reasonable interpretation to the claim limitation, wherein a vector generator would have been an obvious variant of scoring sentences with respect to both statistical and linguistic features such that a centroid query vector is calculated using high frequency document words and the title of the document wherein each sentence is scored accordingly (e.g. using Score formula on page 122) and then ordered in a summary according to rank order, to a person of ordinary skill in the art at the time the invention was made.

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Billheimer '825 teaching, wherein creating a weighted document term-frequency vector for said document, to includes a means of creating a weighted document term-frequency vector for said document from a vector generator of Goldstein's teaching. One of ordinary skill in the art would have been motivated to perform such a modification to produce a query-relevant summary base upon the a weighted sentence term-

frequency vector and accordance with relevance to said weighted document (Goldstein at pages 121-122).

In regard to dependent claim 2, incorporate substantially similar subject matter as cited in claim 1 above, and further view of the following, and is similarly rejected along the same rationale,

recreating said weighted document term-frequency vector in accordance with said deleting and said eliminating, however (Goldstein at pages 121-122, discloses the method of generating summaries by text extraction, wherein scoring sentences with respect to both statistical and linguistic features such that a centroid query vector is calculated using high frequency document words and the title of the document wherein each sentence is scored accordingly (e.g. using Score formula on page 122) and then ordered in a summary according to rank order, wherein each sentence is scored according to the following formula and then ordered in a summary according to rank order,

$$Score(S_i) = \lambda \sum_{s \in S} w_s * (Q_s \cdot S_i) + (1 - \lambda) * \sum_{l \in L} w_l * (L_l \cdot S_i)$$

where S is the set of statistical features, L is the set of linguistic features, Q is the query, and w is the weights for the features in that set. These weights can be tuned according to the type of data set used and the type of summary desired. An ideal query-relevant text summary must contain the relevant information to fulfill a user's information seeking goals, as well as eliminates irrelevant and redundant information) Examiner read the above in the broadest reasonable interpretation to the claim limitation, wherein a vector generator would have been an obvious variant of scoring sentences with respect to both statistical and linguistic features such that a

centroid query vector is calculated using high frequency document words and the title of the document wherein each sentence is scored accordingly (e.g. using Score formula on page 122) and then ordered in a summary according to rank order, to a person of ordinary skill in the art at the time the invention was made.

and selectively repeating said computing, said selecting, said deleting, said eliminating, and said recreating, however (Goldstein at pages 121-122, discloses the method of generating summaries by text extraction, wherein each sentence is scored according to the following formula and then ordered in a summary according to rank order,

$$Score(S_i) = \lambda \sum_{s \in S} w_s * (Q_s \cdot S_i) + (1 - \lambda) * \sum_{l \in L} w_l * (L_l \cdot S_i)$$

where S is the set of statistical features, L is the set of linguistic features, Q is the query, and w is the weights for the features in that set. These weights can be tuned according to the type of data set used and the type of summary desired, wherein scoring sentences with respect to both statistical and linguistic features such that a centroid query vector is calculated using high frequency document words and the title of the document wherein each sentence is scored accordingly (e.g. using Score formula on page 122). The evaluators also marked each document as relevant or not relevant to the topic and selected the three most relevant sentences for each article from the sentences that they had marked relevant (yielding a most relevant sentence data set of 1-9 sentences per document). This set has an average of 5.6 sentences per document and 58.2% of the relevant sentence summaries contain the first sentence. An ideal query-relevant text summary must contain the relevant information to fulfill a user's information seeking goals, as well as eliminates irrelevant and redundant information).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Billheimer '825 teaching, wherein creating a weighted document term-frequency vector for said document, to includes a means of recreating said weighted document term-frequency vector in accordance with said deleting and said eliminating, and selectively repeating said computing, said selecting, said deleting, said eliminating, and said recreating, of Goldstein's teaching. One of ordinary skill in the art would have been motivated to perform such a modification to produce a query-relevant summary base upon the a weighted sentence term-frequency vector and accordance with relevance to said weighted document (Goldstein at pages 121-122).

In regard to dependent claim 3, selectively repeating is terminated when a predetermined number of sentences has been selected, however (Goldstein at pages 121-122, discloses the method of generating summaries by text extraction, wherein each sentence is scored and then ordered in a summary according to rank order. Also include the evaluators that marked each document as relevant or not relevant to the topic and selected the three most relevant sentences for each article from the sentences that they had marked relevant (yielding a most relevant sentence data set of 1-9 sentences per document). This set has an average of 5.6 sentences per document and 58.2% of the relevant sentence summaries contain the first sentence. An ideal query-relevant text summary must contain the relevant information to fulfill a user's information seeking goals, as well as eliminates irrelevant and redundant information).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Billheimer '825 teaching, wherein creating a weighted document term-frequency vector for said document, to includes a means of selectively repeating

is terminated when a predetermined number of sentences has been selected of Goldstein's teaching. One of ordinary skill in the art would have been motivated to perform such a modification to produce a query-relevant summary base upon the a weighted sentence term-frequency vector and accordance with relevance to said weighted document (Goldstein at pages 121-122).

In regard to dependent claim 4, calculating an inner product of said weighted sentence term-frequency vector and said weighted document term-frequency vector however (Goldstein at pages 121-122, discloses the method of generating summaries by text extraction, wherein each sentence is scored according to the following formula and then ordered in a summary according to rank order,

$$Score(S_i) = \lambda \sum_{s \in S} w_s * (Q_s \cdot S_i) + (1 - \lambda) * \sum_{l \in L} w_l * (L_l \cdot S_i)$$

where S is the set of statistical features, L is the set of linguistic features, Q is the query, and w is the weights for the features in that set. These weights can be tuned according to the type of data set used and the type of summary desired, wherein scoring sentences with respect to both statistical and linguistic features such that a centroid query vector is calculated using high frequency document words and the title of the document wherein each sentence is scored accordingly (e.g. using Score formula on page 122). The evaluators also marked each document as relevant or not relevant to the topic and selected the three most relevant sentences for each article from the sentences that they had marked relevant (yielding a most relevant sentence data set of 1-9 sentences per document). This set has an average of 5.6 sentences per document and 58.2% of the relevant sentence summaries contain the first sentence. An ideal query-relevant text

summary must contain the relevant information to fulfill a user's information seeking goals, as well as eliminates irrelevant and redundant information) Examiner read the above in the broadest reasonable interpretation to the claim limitation, wherein the inner product of a weighted sentence term-frequency vector and relevance to said weighted document would have been an obvious variant of calculation of a centroid query vector such as

$$Score(S_i) = \lambda \sum_{s \in S} w_s * (Q_s \cdot S_i) + (1 - \lambda) * \sum_{l \in L} w_l * (L_l \cdot S_i)$$

since inner product is well known mathematical method of an operation on two vectors, which produces a scalar and then ordered in a summary according to rank order, to a person of ordinary skill in the art at the time the invention was made.

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Billheimer '825 teaching, wherein creating a weighted document term-frequency vector for said document, to includes a means of calculating an inner product of said weighted sentence term-frequency vector and said weighted document term-frequency vector of Goldstein's teaching. One of ordinary skill in the art would have been motivated to perform such a modification to produce a query-relevant summary base upon the a weighted sentence term-frequency vector and accordance with relevance to said weighted document (Goldstein at pages 121-122).

In regard to dependent claim 5, wherein said creating a weighted sentence term-frequency vector comprises implementing a local weighting function and implementing a global weighting function, however (Goldstein at pages 121-122, discloses the method of generating summaries by text extraction, wherein scoring sentences with respect to both

statistical and linguistic features such that a centroid query vector is calculated using high frequency document words and the title of the document wherein each sentence is scored accordingly (e.g. using Score formula on page 122) and then ordered in a summary according to rank order) Examiner read the above in the broadest reasonable interpretation to the claim limitation, wherein implementing a local weighting function and implementing a global weighting function would have been an obvious variant of scoring sentences with respect to both statistical and linguistic features such that a centroid query vector is calculated using high frequency document words and the title of the document wherein each sentence is scored accordingly (e.g. using Score formula on page 122) and then ordered in a summary according to rank order, to a person of ordinary skill in the art at the time the invention was made.

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Billheimer '825 teaching, wherein creating a weighted document term-frequency vector for said document, to includes a means of creating a weighted sentence term-frequency vector comprises implementing a local weighting function and implementing a global weighting function of Goldstein's teaching. One of ordinary skill in the art would have been motivated to perform such a modification to produce a query-relevant summary base upon the a weighted sentence term-frequency vector and accordance with relevance to said weighted document (Goldstein at pages 121-122).

In regard to dependent claim 6, incorporate substantially similar subject matter as cited in claim 4 above, and further view of the following, and is similarly rejected along the same rationale,

normalizing each said weighted sentence term-frequency vector, however (Goldstein at pages 121-125, discloses the method of generating summaries by text extraction, wherein analysis of news-article summaries generated by sentence selection. Sentences are ranked for potential inclusion in the summary using a weighted combination of statistical and linguistic features. The statistical features were adapted from standard IR methods. The potential linguistic ones were derived from an analysis of news-wire summaries. To evaluate these features we use a normalized version of precision-recall curves, with a baseline of random sentence selection (see the normalized version as formulas (1) ans (2) on page 125 for details).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Billheimer '825 teaching, wherein creating a weighted document term-frequency vector for said document, to includes a means of normalizing each said weighted sentence term-frequency vector of Goldstein's teaching. One of ordinary skill in the art would have been motivated to perform such a modification to produce a query-relevant summary base upon the a weighted sentence term-frequency vector and accordance with relevance to said weighted document (Goldstein at pages 121-122).

In regard to dependent claim 7, creating a weighted document term-frequency vector comprises implementing a local weighting function and implementing a global weighting function (Billheimer '825 at col. 4, line 35 through col. 6, line 40, discloses an information retrieval method, wherein term and document visualization, term and document clustering, term and document classification, summarization of individual documents and groups of documents, and document cross-referencing. This is accomplished by representing the text of a document collection using subspace transformations. This subspace transformation

representation is performed by: constructing a term frequency matrix of the term frequencies for each of the documents, transforming the term frequencies for statistical purposes, and projecting the documents or the terms into a lower dimensional subspace) Examiner read the above in the broadest reasonable interpretation to the claim limitation, wherein implementing a local weighting function and implementing a global weighting function would have been an obvious variant of constructing a term frequency matrix of the term frequencies for each of the documents, transforming the term frequencies for statistical purposes, and projecting the documents or the terms into a lower dimensional subspace, to a person of ordinary skill in the art at the time the invention was made.

In regard to dependent claim 8, creating a weighted document term-frequency vector comprises normalizing each said weighted document term-frequency vector, however (Goldstein at pages 121-125, discloses the method of generating summaries by text extraction, wherein analysis of news-article summaries generated by sentence selection. Sentences are ranked for potential inclusion in the summary using a weighted combination of statistical and linguistic features. The statistical features were adapted from standard IR methods. The potential linguistic ones were derived from an analysis of news-wire summaries. To evaluate these features we use a normalized version of precision-recall curves, with a baseline of random sentence selection (see the normalized version as formulas (1) ans (2) on page 125 for details).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Billheimer '825 teaching, wherein creating a weighted document term-frequency vector for said document, to includes a means of normalizing each said weighted sentence term-frequency vector of Goldstein's teaching. One of ordinary skill in

the art would have been motivated to perform such a modification to produce a query-relevant summary base upon the a weighted sentence term-frequency vector and accordance with relevance to said weighted document (Goldstein at pages 121-122).

In regard to dependent claims 10, is directed to a computer system for performing the method of claim 2 and is similarly rejected along the same rationale.

In regard to dependent claim 11, is directed to a computer system for performing the method of claims 2-3 and is similarly rejected along the same rationale.

In regard to dependent claim 12, is directed to a computer system for performing the method of claim 3 and is similarly rejected along the same rationale.

In regard to independent claim 13, incorporate substantially similar subject matter as cited in claims 1-3 above, and is similarly rejected along the same rationale.

In regard to dependent claim 14, incorporate substantially similar subject matter as cited in claim 2 above, and is similarly rejected along the same rationale.

In regard to dependent claim 15, incorporate substantially similar subject matter as cited in claim 3 above, and is similarly rejected along the same rationale.

In regard to dependent claim 16, incorporate substantially similar subject matter as cited in claim 4 above, and is similarly rejected along the same rationale.

In regard to dependent claim 17, incorporate substantially similar subject matter as cited in claim 5 above, and is similarly rejected along the same rationale.

In regard to dependent claim 18, incorporate substantially similar subject matter as cited in claim 6 above, and is similarly rejected along the same rationale.

In regard to dependent claim 19, incorporate substantially similar subject matter as cited in claim 7 above, and is similarly rejected along the same rationale.

In regard to dependent claim 20, incorporate substantially similar subject matter as cited in claim 8 above, and is similarly rejected along the same rationale.

5. **Claims 21-32** are rejected under 35 U.S.C. 103(a) as being unpatentable by Billheimer et al. US Patent No. 6,611,825 B1 - filed 06/09/1999 (hereinafter Billheimer '825), in view of Goldstein et al. " Summarizing Text Documents: Sentence Selection and Evaluation Metrics" published 08/1999 by ACM Press (hereinafter Goldstein), further in view of Furnas et al. "Information Retrieval using a Singular Value Decomposition Model of Latent Semantic Structure" published 1988 by ACM Press (hereinafter Furnas) .

In regard to independent claim 21, incorporate substantially similar subject matter as cited in claim 1 above, and further view of the following and is similarly rejected along the same rationale,

Billheimer '825 does not explicitly teach, **constructing a terms-by-sentences matrix for said document, however** (Goldstein at pages 121-127, discloses the method of Evaluation Metrics of summary evaluation, wherein sentences were uses as the underlying unit and evaluated summarization although a list of words, an index or table of content indicate relevance. Further discloses the normalized version as shows in equations (1) and (2) and further discloses at page 126, a summarizer uses the cosine distance metric (of the SMART search engine) to score sentences with respect to a query. For query-relevant summaries, the query is constructed from terms,

each sentence in said document is represented by a column vector of a transpose of said right singular vector matrix; ranking each right singular vector in said right singular vector matrix (Goldstein at pages 121-127, discloses the method of Evaluation Metrics of summary evaluation, wherein sentences were uses as the underlying unit and evaluated summarization although a list of words, an index or table of content indicate relevance. Further discloses the normalized version as shows in equations (1) and (2) and further discloses at page 126, a summarizer uses the cosine distance metric (of the SMART search engine) to score sentences with respect to a query. For query-relevant summaries, the query is constructed from terms.

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Billheimer '825 teaching, wherein creating a weighted document term-frequency vector for said document, to includes a means of constructing a terms-by-sentences matrix for said document and each sentence in said document is represented by a column vector of a transpose of said right singular vector matrix; ranking each right singular vector in said right singular vector matrix of Goldstein's teaching. One of ordinary skill in the art would have been motivated to perform such a modification to produce a query-relevant summary base upon the a weighted sentence term-frequency vector and accordance with relevance to said weighted document (Goldstein at pages 121-122).

Billheimer '825 and Goldstein do not explicitly teach, **performing singular value decomposition on said terms-by-sentences matrix to obtain a singular value matrix, and a right singular vector matrix**, however (Furnas at pages 465-480, discloses the method of automatic indexing and retrieval of the association of terms with documents and the detection of

relevant documents on the basis of terms found in queries using Singular-value decomposition, wherein queries are represented as pseudo-documents vectors formed from weighted combinations of terms, and documents are ordered by their similarity to the query. In a standard vector-based version of term matching, the similarity of two documents is obtained by comparing, e.g., using an inner product or cosine measure, the corresponding two column vectors of the raw data matrix X. A query is represented as a sort of pseudo document, i.e., a column vector of term frequencies Xeq, which is similarly compared against columns of X, and the best matches found, also Furnas at pages 469-472, also see Fig. 3-4, discloses in detail the SVD Model of singular value decomposition of X see equation (1) page 469, wherein Tm and Dm ate matrices of left and right singular vector and Sm is diagonal matrix of singular values m,

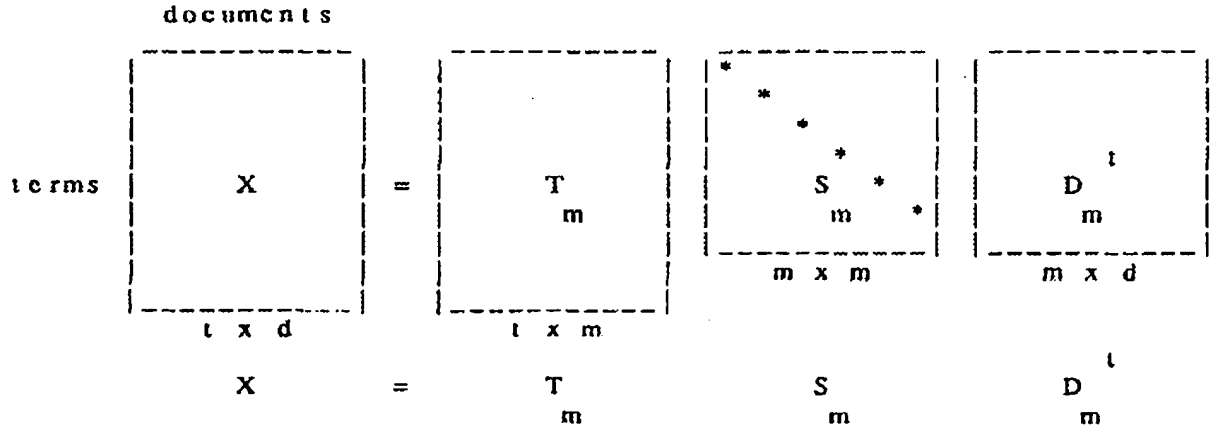

Figure 3. Singular value decomposition of the term x document matrix, **X**. Where :
$\mathbf{T}_m$ has orthogonal, unit-length columns ($\mathbf{T}_m{}^t\mathbf{T}_m = \mathbf{I}$)
$\mathbf{D}_m$ has orthogonal, unit-length columns ($\mathbf{D}_m{}^t\mathbf{D}_m = \mathbf{I}$)
$\mathbf{S}_m$ is the diagonal matrix of singular values
t is the number of rows of $\mathbf{X}$
d is the number of columns of $\mathbf{X}$
m is the rank of $\mathbf{X}$ ($\leq \min(t, d)$)

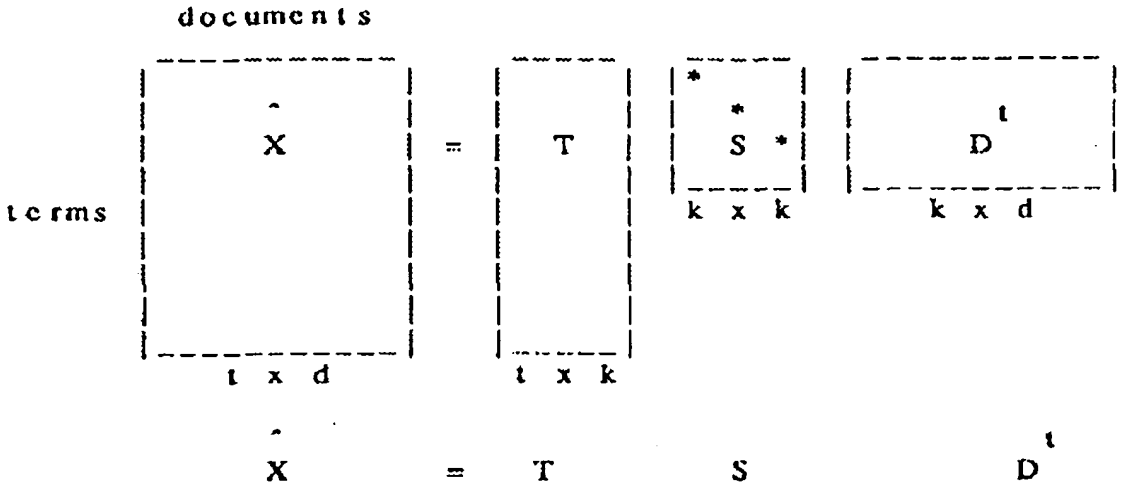

Figure 4. Reduced singular value decomposition of the term x document matrix, $\mathbf{X}$. Notation is as in the previous figure except that k ($\leq m$) is the chosen number of dimensions (factors) in the reduced model.

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Billheimer '825 teaching, wherein creating a weighted document term-frequency vector for said document, to includes a means of constructing a terms-by-sentences matrix for said document and each sentence in said document is represented by a column vector of a transpose of said right singular vector matrix; ranking each right singular vector in said right singular vector matrix of Goldstein's teaching, further to include a means of performing singular value decomposition on said terms-by-sentences matrix to obtain a singular value matrix, and a right singular vector matrix of Furnas' teaching. One of ordinary skill in the art would have been motivated to perform such a modification to produce a query-relevant summary base upon the a weighted sentence term-frequency vector and accordance with relevance to said weighted document (Goldstein at pages 121-122).

In regard to dependent claim 22, repeating said selecting (as taught by Billheimer '825 at col. 9, line 50 through col. 10, line 5).

In regard to dependent claim 23, repeating is terminated when a predetermined number of sentences has been selected" (as taught by Billheimer '825 at col. 9, line 50 through col. 10, line 5).

In regard to dependent claim 24, incorporate substantially similar subject matter as cited in claims 1 and 21 above, and is similarly rejected along the same rationale.

In regard to dependent claim 25, incorporate substantially similar subject matter as cited in claims 1, 7 and 21 above, and is similarly rejected along the same rationale.

In regard to independent claim 26, incorporate substantially similar subject matter as cited in claims 1, 21 and is similarly rejected along the same rationale.

In regard to dependent claims 27-28 consecutively, incorporate substantially similar subject matter as cited in claims 22-23 consecutively and are similarly rejected along the same rationale.

In regard to independent claim 29, incorporate substantially similar subject matter as cited in claims 13, 21 above, and is similarly rejected along the same rationale.

In regard to dependent claims 30-32 consecutively, incorporate substantially similar subject matter as cited in claims 23-25 consecutively, and are similarly rejected along the same rationale.

Response to Arguments

6. Applicant's Remarks filed 04/11/2006 have been fully considered but they are not persuasive. The reason is set forth in the current Office Action cited above and further view of the following:

Brief description of cited prior arts:

Billheimer et al discloses a methodology of summarizing a document performed by constructing a term frequency matrix of the term frequencies for each document (see Billheimer at the Abstract).

Goldsteins et al discloses a text-span extraction and ranking using the methodology that assigns weighted scores for both statically and linguistic features in the text span (see Goldstein page 21 right col. last paragraph).

Furnas et al discloses an Information Retrieval using a Singular Value Decomposition Model of Latent Semantic Structure, wherein Singular-value decomposition is used to decompose a large term by document matrix into orthogonal factors from which the original matrix can be approximated by linear combination; both documents and terms are represented as vectors in a dimensional space, (see Furnas at page 465 the Abstract).

Response to Arguments:

Beginning on page 2 of the Remarks (hereinafter the remarks), Applicant argues the following issues, which are accordingly addressed below.

Applicant's arguments, on pages 2-3 of the remarks that Billheimer in combination with Goldstein fail to teach, for each sentence in said document, creating a weighted sentence term-frequency vector.

The examiner respectfully disagrees. The examiner respectfully notes that Billheimer et al discloses a methodology of summarizing a document performed by constructing a term frequency matrix of the term frequencies for each document (see Billheimer at the Abstract); but

Billheimer does not explicitly teach, for each sentence in said document, creating a weighted sentence term-frequency vector. However, as cited in the rejection above and further Goldsteins et al discloses a text-span extraction and ranking using the methodology that assigns weighted scores for both statically and linguistic features in the text span (see Goldstein page 21 right col. last paragraph). Also Goldstein teaches using Score formula on page 122 and then ordered in a summary according to rank order, the summarization of the document sometimes called Abstraction, produces a fixed-length generic summary that reflects the key points which the abstractor deems important (see Goldstein page 122 section 2 "Generating Summaries be text Extraction"), Since the resulting summary of a document comes from the a term frequency matrix of the term frequencies for each document (i.e. score, weighted, ranking matrix of term frequency is well known as term frequency vector, to a person of ordinary skill in the art at the time the invention was made), in order to determine the level if relevant and/or important of any particular term in a document for extraditions the text to form a document summary (Abstract).

Therefor the Examiner respectfully maintains the rejection of claims 1-20 for at least the reason cited above at this time.

Applicant's arguments, on pages 3-4 of the remarks that Billheimer in combination with Goldstein and Furnas fail to disclose or suggest all of the claim limitation.

The examiner respectfully disagrees. The examiner respectfully notes that Billheimer and Goldstein teach all the limitation, but "using singular value decomposition" (see responses to the remarks of claims 1-20 cites above). However using the broadest interpretation, to cure the deficiencies of Billheimer and Goldstein the examiner respectfully notes that Furnas et al discloses an Information Retrieval using a Singular Value Decomposition Model of

Latent Semantic Structure, wherein Singular-value decomposition is used to decompose a large term by document matrix into orthogonal factors from which the original matrix can be approximated by linear combination; both documents and terms are represented as vectors in a dimensional space, (see Furnas at page 465 the Abstract).

It would have been obvious to one of ordinary skill in the art at the time of the invention to apply Furnas's Singular Value Decomposition into Billheimer and Goldstein teaching to decompose a large term by document matrix into orthogonal factors from which the original matrix can be approximated by linear combination; both documents and terms are represented as vectors in a dimensional space, (see Furnas at page 465 the Abstract).

Therefor the Examiner respectfully maintains the rejection of claims 21-32 for at least the reason cited above at this time.

Conclusion

7. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event,

however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Quoc A. Tran whose telephone number is (571) 272-4103. The examiner can normally be reached on Monday through Friday from 9 AM to 5 PM EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Herndon R. Heather can be reached on (571) -272-4136. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Quoc A, Tran
Patent Examiner
Technology Center 2176
January 5, 2006

William L. Bashore
WILLIAM BASHORE
PRIMARY EXAMINER